Markov Processes from K. Itô's Perspective

Daniel W. Stroock

ANNALS OF MATHEMATICS STUDIES

Markov Processes from K. Itô's Perspective

by

Daniel W. Stroock

PRINCETON UNIVERSITY PRESS

PRINCETON AND OXFORD

2003

Published by Princeton University Press, 41 William Street,
Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press,
3 Market Place, Woodstock, Oxfordshire OX20 1SY

The Annals of Mathematics Studies are edited by
John N. Mather and Elias M. Stein

ISBN 0-691-11542-7 (cloth)
ISBN 0-691-11543-5 (pbk.)

British Library Cataloging-in-Publication Data is available.

The publisher would like to acknowledge the author of this
volume for providing the camera-ready copy from which this book was printed

Printed on acid-free paper ∞
www.pup.princeton.edu

Printed in the United States of America

1 3 5 7 9 10 8 6 4 2

1 3 5 7 9 10 8 6 4 2
(Pbk.)

Annals of Mathematics Studies

Number 155

This book is dedicated to the only person to whom it could be:

Kiyosi Itô

Preface

In spite of (or, maybe, because of) his having devoted much of his life to the study of probability theory, Kiyosi Itô is not a man to leave anything to chance. Thus, when Springer-Verlag asked S.R.S. Varadhan and me to edit a volume [14] of selected papers by him, Itô wanted to make sure that we would *get it right*. In particular, when he learned that I was to be the author of the introduction to that volume, Itô , who had spent the preceding academic year at the University of Minnesota, decided to interrupt his return to Japan with a stop at my summer place in Colorado so that he could spend a week tutoring me. In general, preparing a volume of selected papers is a pretty thankless task, but the opportunity to spend a week being tutored by Itô more than compensated for whatever drudgery the preparation of his volume cost me. In fact, that tutorial is the origin of this book.

Before turning to an explanation of what the present book contains, I cannot, and will not, resist the temptation to tell my readers a little more about Itô the man. Specifically, I hope that the following, somewhat trivial, anecdote from his visit to me in Colorado will help to convey a sense of Itô's enormous curiosity and his determination to understand the world in which he lives. On a day about midway through the week which Itô spent in Colorado, I informed him that one of my horses was to be shod on the next day and that, as a consequence, it might be best if we planned to suspend my tutorial for a day. Considering that it was he who had taken the trouble to visit me, I was somewhat embarrassed about asking Itô to waste a day. However, Itô's response was immediate and completely positive. He wanted to learn how we Americans put shoes on our horses and asked if I would pick him up at his hotel in time for him to watch. I have a vivid memory of Itô bending over right next to the farrier, all the time asking a stream of questions about the details of the procedure. It would not surprise me to learn that, after returning home, Itô not only explained what he had seen but also suggested a cleverer way of doing it. Nor would it surprise me to learn that the farrier has avoided such interrogation ever since.

The relevance of this anecdote is that it highlights a characteristic of Itô from which I, and all other probabilists in my generation, have profited immeasurably. Namely, Itô's incurable pedagogic itch. No matter what the topic, Itô is driven to master it in a way that enables him to share his insights with the rest of us. Certainly the most renowned example of Itô's skill is his introduction of stochastic differential equations to explain the Kolmogorov–Feller theory of Markov processes. However, the virtues of his theory were not immediately recognized. In fact, sometime during the week of my tutorial, Itô confided in me his disappointment with the initial reception which his theory received. It seems that J.L. Doob was the first person to fully appreciate what Itô had done. To wit, Doob not only played a crucial role in arranging for Itô's memoirs [11] to be published by the AMS, he devoted §5 of the ninth chatper in his own famous book [6] to explaining, extending, and improving Itô's theory. However, Doob's book is sufficiently challenging that hardly anyone who made it to §5 of Chapter IX was ready to absorb its content. Thus, Itô's theory did not receive the attention that it deserves until H.P. Mc Kean, Jr., published his lovely little book [23].

Like Lerey's theory of spectral sequences, Itô's theory was developed as a pedagogic device, and, like Leray's theory, Itô's theory eventually took on a life of its own. Both Itô's and Mc Kean's treatments of the theory concentrate on practical matters: the application of stochastic integration to the path-wise construction of Markov processes. However, around the same time that Mc Kean's book appeared, P.A. Meyer and his school completed the program which Doob had initiated in Chapter IX. In particular, Meyer realized that his own extension of Doob's decomposition theorem allowed him to carry out for essentially general martingales the program which Doob had begun in his book, and, as a result, Itô's theory became an indispensable tool for the Strasbourg School of Probability Theory. Finally, for those who found the Strasbourg version too arcane, the theory was brought back to earth in the beautiful article [21], where Kunita and Watanabe showed that, if one takes the best ideas out of the Strasbourg version and tempers them with a little reality, the resulting theory is easy, æsthetically pleasing, and remarkably useful.

In the years since the publication of Mc Kean's book, there have been lots of books written about various aspects of Itô's theory of stochastic integration. Among the most mathematically influential of these are the book by Delacherei and P.A. Meyer's [5], which delves more deeply than most will care into the intimate secrets of stochastic processes, N. Ikeda and S. Watanabe's [16], which is remarkable for both the breadth and depth of the material covered, and the book by D. Revuz and M. Yor's [27], which demonstrates the power of stochastic integration by showing how it can be used to elucidate the contents Itô's book with Mc Kean [15], which, ironically, itself contains *no* stochastic integration. Besides these, there are several excellent, more instruction oriented texts: the ones by B. Øksendal [25] and by K.L. Chung and R. Williams [3] each has devoted followers. In addition, the lure of mammon has encouraged several authors to explain stochastic integration to economists and to explain economics to stochastic integrationists.[1]

In view of the number of books which already exist on the subject, one can ask, with reason, whether yet another book about stochastic integration is really needed. In particular, what is my excuse for writing this one? My answer is that, whether or not it is needed, I have written this book to redress a distortion which has resulted from the success of Itô's theory of stochastic integration. Namely, Itô's stochastic integration theory is a *secondary theory*, the *primary theory* being the one which grew out of Itô's ideas about the structure of Markov processes. Because his primary theory is the topic of Chapters 1 through 4, I will restrict myself here to a few superficial comments. Namely, as is explained in the introduction to [14], when, as a student at Tokyo University, Itô was assigned the task of explaining the theory of Markov processes to his peers, he had the insight that Kolmogorov's equations arise naturally if one thinks of a Markov process as *the integral curve of a vector field on the space* $\mathbf{M}_1(\mathbb{R}^n)$ *of probability measures on* $\mathbb{R}^n$. Of course, in order to think about vector fields on $\mathbf{M}_1(\mathbb{R}^n)$, one has to specify a coordinate system, and Itô realized that the one which not only leads to Kolmogorov's equations but also explains their connection to infinitely divisible laws is the coordinate system determined by integrating $\mu \in \mathbf{M}_1(\mathbb{R}^n)$ against smooth functions. When these coordinates are adopted, the infinitely divisible flows play the role of "rays," and Kolmogorov's equations arise as the equations which determine the integral curve of a vector field composed of these rays.[2]

From a conventional standpoint, Chapter 1 is a somewhat peculiar introduction to the theory of continuous-time Markov processes on a finite state space. Namely, because it is the setting in which all technical difficulties disappear, Chapter 1 is devoted to the development of Itô's ideas in the setting of the n-point space $\mathbb{Z}_n \equiv \{0, \dots, n-1\}$. To get started, I first give $\mathbf{M}_1(\mathbb{Z}_n)$ the differentiable structure that it inherits as a simplex in $\mathbb{R}^n$. I then think of $\mathbb{Z}_n$ as an

---

[1] Whatever the other economic benefits of this exercise have been, it certainly provided (at least for a while) a new niche in the job market for a generation of mathematicians.

[2] It should be mentioned that, although this perspective is enormously successful, it also accounts for one of the most serious shortcomings of Itô's theory. Namely, Itô's ideas apply only to Markov processes which are smooth when viewed in the coordinate system which he chose. Thus, for example, they cannot handle diffusions corresponding to divergence-form, strictly elliptic operators with rough coefficients, even though the distribution of such a diffusion is known to be, in many ways, remarkably like that of Brownian motion.

Abelian group (under addition modulo n), develop the corresponding theory of infinitely divisible laws, and show that the inherited differentiable structure has a natural description in terms of infinitely divisible flows, which now play the role that rays play in Euclidean differential geometry. Having put a differentiable structure on $\mathbf{M}_1(\mathbb{Z}_n)$, one knows what vector fields are there, and so I can discuss their integral curves. Because it most clearly highlights the analogy between the situation here and the one which is encountered in the classical theory of ordinary differential equations, the procedure which I use to integrate vector fields on $\mathbf{M}_1(\mathbb{Z}_n)$ is the Euler scheme: the one in which the integral curve is constructed as the limit of "polygonal" curves. Of course, in this setting, "polygonal" means "piecewise infinitely divisible." In any case, the integral curve of a general vector field on $\mathbf{M}_1(\mathbb{Z}_n)$ gives rise to a nonlinear Markov flow: one for which the transition mechanism depends not only on where the process presently is but also on what its present distribution is. In order to get usual Markov flows (i.e., linear ones with nice transition probability functions), it is necessary to restrict one's attention to vector fields which are affine with respect to the obvious convex structure of $\mathbf{M}_1(\mathbb{Z}_n)$, at which point one recovers the classical structure in terms of Kolmogorov's forward and backward equations.

The second half of Chapter 1 moves the considerations of the first half to a pathspace setting. That is, in the second half I show that the flows produced by integrating vector fields on $\mathbf{M}_1(\mathbb{Z}_n)$ can be realized in a canonical way as the one-time marginal distributions of a probability measure $\mathbb{P}$ on the space $D\big([0,\infty);\mathbb{Z}_n\big)$ of right-continuous paths, piecewise constant $p:[0,\infty)\longrightarrow\mathbb{Z}_n$. After outlining Kolmogorov's completely general approach to the construction of such measures, the procedure which I adopt and emphasize here is Itô's. Namely, I begin by constructing the $\mathbb{P}$, known as the Lévy process, which corresponds to an infinitely divisible flow. I then show how the paths of a Lévy process can be massaged into the paths of a general Markov process on $\mathbb{Z}_n$. More precisely, given a transition probability function on $\mathbb{Z}_n$, I show how to produce a mapping, the "Itô map," of $D\big([0,\infty);\mathbb{Z}_n\big)$ into itself in such a way that the $\mathbb{P}$ corresponding to the given transition probability function is the image under the Itô map of an appropriate Lévy process. The chapter ends by showing that the construction of the Itô map can be accomplished by the pathspace analog of the Euler approximation scheme used earlier to integrate vector fields on $\mathbf{M}_1(\mathbb{Z}_n)$.

The second and third chapters recycle the ideas introduced in Chapter 1, only this time when the state space is $\mathbb{R}^n$ instead of $\mathbb{Z}_n$. Thus, Chapter 2 develops the requisite machinery, especially the Lévy–Khinchine formula,[3] about

[3] In the hope that it will help to explain why these techniques apply to elliptic, parabolic,

infinitely divisible flows, and then shows[4] that, once again, the infinitely divisible flows are the rays in this differentiable structure. Once this structure has been put in place, I integrate vector fields on $\mathbf{M}_1(\mathbb{R}^n)$ using Euler's approximation scheme and, after specializing to the affine case, show how this leads back to Kolmogorov's description of Markov transition probability functions via his forward and backward equations.

Before I can proceed to the pathspace setting, I have to develop the basic theory of Lévy processes on $\mathbb{R}^n$. That is, I have to construct the pathspace measures corresponding to infinitely divisible flows in $\mathbf{M}_1(\mathbb{R}^n)$, and, for the sake of completeness, I carry this out in some detail in the last part of Chapter 2. Having laid the groundwork in Chapter 2, I introduce in Chapter 3 Itô's procedure for converting Lévy processes into more general Markov processes. Although the basic ideas are the same as those in the second half of Chapter 1, everything is more technically complicated here and demands greater care. In fact, it does not seem possible to carry out Itô's procedure in complete generality, and so Chapter 3 includes an inquiry into the circumstances in which his procedure does work. Finally, Chapter 3 concludes with a discussion of some examples which display both the virtues and the potential misinterpretation of Itô's theory.

Chapter 4 treats a number of matters which are connected with Itô's construction. In particular, it is shown that, under suitable conditions, his construction yields measures on pathspace which vary smoothly as a function of the starting point, and this observation leads to a statement of uniqueness which demonstrates that the pathspace measures at which he arrives are canonically connected to the affine vector fields from which they arise.

In some sense, Chapter 4 could, and maybe should, be the end of this book. Indeed, by the end of Chapter 4, the essence of Itô's theory has been introduced, and all that remains is to convince afficionados of Itô's theory of stochastic integration that the present book is about their subject. Thus, Chapter 5 develops Itô's theory of stochastic integration for Brownian motion and applies it to the constructions in Chapter 3 when Brownian motion is the underlying Lévy process. Of course, Itô's famous formula is the centerpiece of all this. Chapter 6 showcases several applications of Itô's stochastic integral theory, especially his formula. Included are Tanaka's treatment of local time, the Cameron–Martin formula and Girsanov's variation thereof, pinned Brownian motion, and

---

but not hyperbolic equations, I have chosen to base my proof of their famous formula on the minimum principle.

[4] Unfortunately, certain technical difficulties prevented me from making the the connection here as airtight as it is in the case of $\mathbf{M}_1(\mathbb{Z}_n)$.

Itô's treatment of Wiener's theory of homogeneous chaos.

Kunita and Watanabe's extension of Itô's theory to semimartingales is introduced in Chapter 7 and is used there to prove various representation theorems involving continuous martingales. Finally, Chapter 8 deals with Stratonovich's variant on Itô's integration theory. The approach which I have taken to this topic emphasizes the interpretation of Stratonovich's integral, as opposed to Itô's , as the one at which someone schooled in L. Schwartz's distribution theory would have arrived. In particular, I stress the fact that Stratonovich's integral enjoys coordinate inveriance properties which, as is most dramatically demonstrated by Itô's formula, Itô's integral does not. In any case, once I have introduced the basic ideas, the treatment here follows closely the one developed in [38]. Finally, the second half of this chapter is devoted to a proof of the "support theorem" (cf. [39] and [40]) for degenerate diffusions along the lines suggested in that article. Although applications to differential geometry are among the most compelling reasons for introducing Stratonovich's theory, I have chosen to not develop them in this book, partly in the hope that those who are interested will take a look at [37].

Before closing, I should say a few words about possible ways in which this book might be read. Of course, I hope that it will eventually be read by at least one person, besides myself, from cover to cover. However, I suspect that most readers will not to do so. In particular, because this book is really an introduction to continuous time stochastic processes, the reader who is looking for the most efficient way to learn how to do stochastic integration (or price an option) is going to be annoyed by Chapters 1 and 2. In fact, a reader who is already comfortable with Brownian motion and is seeking a no frills introduction to the most frequently used aspects of Itô's theory should probably start with Chapter 5 and dip into the earlier parts of the book only as needed for notation. On the other hand, for someone who already knows the nuts and bolts of stochastic integration theory and is looking to acquire a little "culture," a reasonable selection of material would be the contents of Chapter 1, the first half of Chapter 2, and Chapter 3. In addition, all readers may find some topics of interest in the later chapters.

Whatever is the course which they choose, I would be gratified to learn that some of my readers have derived as much pleasure out of reading this book as I have had writing it.

Daniel W. Stroock, August 2002

Markov Processess from K. Itô's Perspective

Contents

CHAPTER 1

Finite State Space, a Trial Run

In his famous article [19], Kolmogorov based his theory of Markov processes on what became known as Kolmogorov's *forward*[1] and *backward* equations. In an attempt to explain Kolmogorov's ideas, K. Itô [14] took a crucial step when he suggested that Kolmogorov's forward equation can be thought of as describing the "flow of a vector field on the space of probability measures." The purpose of this chapter is to develop Itô's suggestion in the particularly easy case when the state space E is finite. The rest of the book is devoted to carrying out the analogous program when $E = \mathbb{R}^n$.

1.1 An Extrinsic Perspective

Suppose that $\mathbb{Z}_n = \{0, \dots, n-1\}$ where $n \geq 1$. Then $\mathbf{M}_1(\mathbb{Z}_n)$ is the set of probability measures $\mu_{\boldsymbol{\theta}} = \sum_{m=0}^{n-1} \theta_m \delta_m$, where each δ_m is the unit point mass at m and $\boldsymbol{\theta}$ is an element of the simplex $\Theta_n \subseteq \mathbb{R}^n$ consisting of vectors whose coordinates are non-negative and add to 1. Furthermore, if Θ_n is given the topology which it inherits as a subset of $\mathbb{R}^n$ and $\mathbf{M}_1(\mathbb{Z}_n)$ is given the topology of weak convergence (i.e., the topology in which convergence is tested in terms of integration against bounded, continuous functions), then the map $\boldsymbol{\theta} \in \Theta_n \longmapsto \mu_{\boldsymbol{\theta}} \in \mathbf{M}_1(\mathbb{Z}_n)$ is a homeomorphism. Thus it is reasonable to attempt using this homeomorphism to gain an understanding of the differentiable structure on $\mathbf{M}_1(\mathbb{Z}_n)$.

§1.1.1. The Structure of Θ_n: It is important to recognize that the differentiable structure which Θ_n inherits as a subset of $\mathbb{R}^n$ is not entirely trivial. Indeed, even when $n = 2$, Θ_n is a submanifold with boundary, and it is far worse when $n \geq 3$. More precisely, its interior $\mathring{\Theta}_n$ is a nice $(n-1)$-dimensional submanifold of $\mathbb{R}^n$, but the boundary $\partial\Theta_n$ of Θ_n breaks into faces which are disconnected submanifolds of dimensions $(n-2)$ through 0. For example,

$$(\theta_1, \theta_2) \in \left\{(\xi_1, \xi_2) \in (0,1)^2 : \xi_1 + \xi_2 < 1\right\} \longmapsto (1 - \theta_1 - \theta_2, \theta_1, \theta_2) \in \mathring{\Theta}_3$$

[1] In the physics literature, Kolmogorov's forward equation is usually called the Fokker-Planck equation.